

**Annex V – Justification of Costs  
Quality of Life – Health Research**

**PROPOSAL NUMBER**  
**PARTNER NAME**

**QLRT-2001-01254**  
**CR 9 University of Leiden, Netherlands**  
**Dr Lodewijk Sandkuijl**

<b>Equipment Item</b>	<b>Technical Description</b>	<b>Amount to be charged to the Project [€]</b>
Equipment name	Description, components; functional description; justification of its needs and tasks where it will be used; and general specifications; possible suppliers, etc.	cost breakdown and total cost to be charged to the project
Computer, Sun Blade 1000	Fast desktop workstation, for carrying out complex statistical analyses, including extra memory, network connections and software	€ 15.000,-
Extra processor for Sun Blade 1000	In the course of the project more and more heavy analyses will be have to be carried out. The advantage of the Sun Blade Series is that it can be extended with a second processor which is even faster than the first processor.	€ 10.000,-
Six desktop PC's	For development and testing of software, and for statistical analyses (extra memory up to 1 Gb is required in order to manage large datasets)	six times € 4.000,-
Notebook computer	For development and testing of distributed software at the various participating centers	€ 4.500,-
Extra hard-disk	Extending capacity of existing network server with 140 Gb	€ 3.000,-
Extra back-up facility	Large datasets need to be backed up on separate tapes	€ 3.500,-
Additional printer	Three network printers are already available, but one additional printer is required for the group of six analysts	€ 2.000,-
<b>TOTAL</b>		€ 62.000,-

<b>Subcontractor or Services</b>	<b>Work description</b>	<b>Total Cost [€]</b>
Company name and country; or t.b.d. <sup>1</sup> . and country; or t.b.d.	Description of the activities/services in general and/or work packages in particular; justification and/or task where this subcontractor will supply its services; previous experience with them; possible subcontractors and when to take the decision	Total costs of the services (100% to be charged to the project)
n.a.	<b>TOTAL</b>	

<b>Consumables</b>	<b>Technical Description</b>	<b>Amount to be charged to the Project [€]</b>
Short description or item #	Description of components and materials; justification of its needs and tasks where it will be used; justification of the calculated values and general features, etc.	cost breakdown and total cost to be charged to the project
n.a.	<b>TOTAL</b>	

<sup>1</sup> T.b.d. : to be detrmined

**Annex V – Justification of Costs  
Quality of Life – Health Research**

<b>Travel</b>	<b>Place</b>	<b>Traveling persons</b>	<b>Purpose and duration</b>	<b>Cost [€]</b>
Number of the month planned	Location planned	number of persons	description and justification. In case of more than 1 person, please, give justifications	Total costs of the travel to be charged to the project
every three months during the first year, frequency in subsequent years t.b.d.	Odense, Oslo, Helsinki, Rome, Amsterdam, Stockholm, Uppsala.	1	see separate justification	€ 20.000,- for first year, € 15.000,- for subsequent years (estimated)
t.b.d., but at least twice during the project, and at least once during the first year	Los Angeles	2	see separate justification, estimated duration three weeks	€ 11.000,-
Scientific meeting	t.b.d.	6	to present the newly developed methods	€ 9.000,-
<b>TOTAL</b>				<b>€ 85.000,-</b>

<b>Other specific project costs</b>	<b>Technical Description</b>	<b>Amount to be charged to the Project [€]</b>
Name	Description of the costs and justification of the needs and tasks where they will be used; justification of the calculated values and general features, etc.	cost breakdown and total cost to be charged to the project
Computing	Use of university mainframe computer	€ 13.000,-
Software licences	Sage (statistical analysis for genetic epidemiology) package, Statexact, STATA, C++, Fortran compiler	€ 8.000,-
<b>TOTAL</b>		<b>€ 21.000,-</b>

Personnel Costs: all details to be provided in CPF form A7.2 1/2  
 IPR costs: all details to be provided in CPF form A7.3 ½

Justification for personnel costs:

The tasks of the statistical core are rather diverse. New methods of analysis and error control will have to be developed, existing methods and software programs will have to be tested, and many large-scale analyses will have to be carried out. A reasonable number of software tools exist for the combined statistical analysis of genotypes and phenotypes, but most of those tools were initially developed for samples of one or a few families. The current project requires analyses on a completely different scale, which will without any doubt lead to considerable logistical problems. Therefore it is imperative that in the early stage of the project all computer programs that we intend to use for the statistical analyses be tested with extremely large simulated data sets in order to detect well in advance any hitherto unknown flaws in the programs or serious shortcomings with respect to performance.

A unique feature of the data that will be analysed in the Genomeutwin project is that they are from different geographical areas in Europe. This allows, or rather requires, that many of the analyses be carried out in a stratified manner. For many of the existing software tools no options have been implemented to carry out such stratified analyses on large datasets. This will require specific modifications.

Altogether the tasks for the statistical analysis core can be divided into four broad groups:

- 1) development of tools for the analysis of longitudinal phenotype data in combination with extensive genotypical data: this will require one post-doc during two-years,
- 2) development and testing of QTL mapping tools, preparation of tools for distributed analysis, and joint analyses of heritability for the selected traits, and localisation of genetic determinants for these traits: this will require one senior statistician for the entire period of the project, one post-doc for two-years, and one PhD student for the entire period of four years,
- 3) development of methodology and tools for analysis of genotype and/or haplotype data in combined dataset in order to detect and describe substructuring (e.g., with respect to linkage disequilibrium) of data according to geographical or other factors: this will require one post-doc for two years and one PhD student for the entire period of four years,
- 4) development of statistical methodology and tools for large-scale association analysis for quantitative traits, including the analysis of specific interaction patterns between SNP genotypes, preparation of tools for distributed analyses, carrying out of combined analyses: this will require one senior statistician for the entire period of four years, and one post-doc for two years.

As correctly pointed out by the reviewers there was an unfortunate error in the original application which led to an obvious discrepancy between the salary versus the number of personnel months (the latter should now become 304 rather than 604).

Justification for travel costs:

As proposed in the original application, frequent visits to the participating centers by the coordinator of the statistical core or one of his statisticians are absolutely essential, particularly in the early stages of the project. In the original application it was suggested that each center be visited once every two months, but it may be possible to reduce the frequency of the visits to once every three months. One of the reasons for such frequent visits which may not be immediately obvious is the extreme urgency for identification and correction of all possible sources of data-errors. It is not frequently realized by those who are not involved in statistical genetics that the impact of data-errors is much more serious in the statistical analysis of family data than in regular statistical analyses. This is due to the statistical dependence that is present in data on related individuals: it has been shown that 1 – 2 % errors is sufficient to make important genetic effects undetectable ! Many error checks will be carried out by the statistical core that are only possible by virtue of comparison within and between the extremely large datasets that will become available for pooled analyses. It is clear that all possible indications of data-errors should be sorted out as soon as possible. It is the experience of the coordinator of the statistical core, Dr L.A. Sandkuijl, during the statistical analyses of more than 100 mapping projects in collaboration with an even larger number of different centers, that direct personal contact is invaluable when sorting out possible data problems. But we will immediately admit that the necessity of the frequent visits will need to be re-evaluated very regularly.

Another not immediately obvious reason for additional travel expenses is the fact that our collaboration is fortunate to count Prof. Kenneth Lange, at UCLA among the members of the statistical core. Prof. Lange is one of the most prominent statistical geneticists in the world (in addition to being a prominent biostatistician), and he is willing to provide consultation for our group with respect to the development of new approaches. This will require, however, several visits to Los Angeles for several of the statisticians in the statistical core to initiate the collaboration. Further contacts can be maintained electronically.

Justification for computing costs:

While the majority of the analyses can be carried out on the workstation and PC's that will be purchased for this project, there will be some analyses that require extremely time-consuming calculations. For such analyses it will be necessary to resort to the university computer center.